Forensic Science International: Genetics 7 (2013) 82–91

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Developmental validation of DogFiler, a novel multiplex for canine DNA profiling in forensic casework☆

Elizabeth Wictum [a,*], Teri Kun [a], Christina Lindquist [a], Julia Malvick [a], Dianne Vankan [b], Benjamin Sacks [c,d]

[a] Forensic Unit, Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis, CA, USA
[b] The Animal Genetics Laboratory, University of Queensland, Australia
[c] Canid Diversity and Conservation Unit, Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis, CA, USA
[d] Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA, USA

ARTICLE INFO

Article history:
Received 12 March 2012
Received in revised form 3 July 2012
Accepted 4 July 2012

Keywords:
Canine
Dog
DogFiler
Non-human DNA
Validation

ABSTRACT

While the analysis of human DNA has been the focus of large-scale collaborative endeavors, non-human forensic DNA analysis has not benefited from the same funding streams and coordination of effort. Consequently, the development of standard marker panels, allelic ladders and allele-specific sequence data comparable to those established for human forensic genetics has lagged. To meet that need for domestic dogs, we investigated sequence data provided by the published 7.6X dog genome for novel short tandem repeat markers that met our criteria for sensitivity, stability, robustness, polymorphic information content, and ease of scoring. Fifteen unlinked tetranucleotide repeat markers were selected from a pool of 3113 candidate markers and assembled with a sex-linked marker into a multiplex capable of generating a full profile with as little as 60 pg of nuclear DNA. An accompanying allelic ladder was assembled and sequenced to obtain detailed repeat motif data. Validation was carried out according to SWGDAM guidelines, and the DogFiler panel has been integrated into forensic casework and accepted in courts across the U.S. Applying various formulae for calculating random match probabilities for inbred populations, estimates for this panel of markers have proven to be comparable to those obtained in human forensic genetics. The DogFiler panel and the associated allelic ladder represent the first published non-human profiling system to fully address all SWGDAM recommendations.

1. Introduction

The molecular analysis of animal DNA is increasingly being admitted into evidence in criminal justice systems around the world. This ranges from crimes against animals such as the poaching of protected wildlife species, animal cruelty and dog fighting, to the human-on-human crimes of rape, robbery and homicide where there is a transfer of animal biological evidence. There are over 70 million pet dogs in the United States with 39% of homes having one or more dogs [1]. Canine DNA in the form of hair, saliva, blood, urine, and feces is abundant in the domestic environment, and consequently is often present on evidence collected during crime-scene investigations. Pet hair in particular has been found to be easily transferrable [2] and is often encountered on the clothing, bedding, bodies, and vehicle interiors of pet owners. While shed hair lacks adherent follicular material needed for nuclear DNA analysis, the grooming behavior of dogs can deposit epithelial cells on the surface of the hair shafts increasing amplification success [3].

The earliest reported use of domestic animal DNA profiling in a criminal case was a 1994 death investigation where cat hairs on a bloody jacket lead investigators back to the victim's estranged common-law husband [4]. Since that time, the forensic analysis of both nuclear and mitochondrial dog DNA has been reported in peer-reviewed journals [5–9]. However, due to the paucity of laboratories routinely performing canine forensic testing, a standard panel of loci has not been established and there is no way to exchange or compare nuclear STR data.

Microsatellite or short tandem repeat (STR) loci have been identified in numerous plant and animal genomes and have become the gold standard for DNA profiling. They are abundant, polymorphic, and easily characterized by polymerase chain reaction (PCR) amplification using fluorescent dye-labeled primers and a laser detection system. Their utility is greatly enhanced when multiple polymorphic markers are combined into a multiplex capable of generating a DNA profile with minimal consumption of resources. Polymorphic canine STR loci have been used extensively for parentage verification, phylogeny, association mapping, linkage

☆ This work was presented at the International Society for Animal Genetics meeting in Amsterdam, The Netherlands, July 2008; and at the American Academy of Forensic Sciences meeting in Denver, CO, USA, February 2009.

* Corresponding author at: Veterinary Genetics Laboratory, Forensic Unit, School of Veterinary Medicine, University of California, One Shields Ave., Davis, CA 95616-8744, USA. Tel.: +1 530 752 2211; fax: +1 530 309 4309.
E-mail address: ejwictum@ucdavis.edu (E. Wictum).

1872-4973/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsigen.2012.07.001

analysis, and conservation genetics [10–18]. Many markers have proven to be acceptable for these purposes and are used routinely by laboratories with access to optimal DNA sources. However, biological evidence in forensic investigations is often limited or degraded and so requires the selection of markers that meet more rigorous criteria. Ideally, autosomal STR markers for forensic testing should be robust, unlinked, highly polymorphic, and contain repeat motifs of four or more base pairs with a low frequency of microvariant alleles. Primer binding sites must be highly conserved, be genus or species-specific, and produce clean profiles with low stutter product [19].

A panel of markers recommended for parentage in domestic dogs by the International Society for Animal Genetics (ISAG) [20] is the only panel standardized across multiple laboratories for canine genotyping. It contains 21 di-nucleotide repeat markers and three tetra-nucleotide repeat markers. The higher stutter product of dinucleotide repeats is an obstacle to mixture deconvolution, so they are less suited for analysis of mixtures than markers with repeat units of four or more nucleotides [21]. Several panels have been reported for use on canine forensic samples [22–25], however, each of those panels contain markers that did not meet our specific criteria due to high mutation rates, lack of robustness, a high frequency of microvariant alleles, overlapping allele ranges, or large allelic ranges that promote peak imbalance and allelic dropout. Furthermore, some panels are amplified in two multiplexes which increases the consumption of template DNA and the costs associated with multiple PCR reactions.

Our goal was to produce a balanced and highly discriminating multiplex optimized for standard PCR conditions that could be implemented in forensic laboratories worldwide. In order to meet that goal, we exploited an opportunity to mine the 7.6X dog genome sequence data (Broad Institute, CanFam2.0) [26] for novel markers that met our criteria. We ultimately selected 15 markers and assembled them into a single multiplex, along with a sex-identification marker [27], capable of generating a full DNA profile with as little as 60 pg of genomic DNA under the reaction and fragment separation conditions given below. This panel—entitled DogFiler—was validated in accordance with the Scientific Working Group for DNA Analysis Methods (SWGDAM) revised guidelines for developmental validation [28] to determine the limitations of the procedure and to assure the accuracy, precision, and reproducibility of test results. An accompanying DNA ladder was created per published recommendations [29] to facilitate the sharing of canine profiles and databases.

2. Materials and methods

2.1. Marker selection

Computational tools were designed to mine the May 2005 assembly of the 7.6X dog genome for novel $GATA_n$ and $GAAA_n$ tetra-nucleotide repeat loci having ten or more repeat units. After masking the published sequence for known microsatellites, we identified 4180 candidate loci in the reference genome maintained in the UCSC *simpleRepeat* database (http://genome.ucsc.edu/cgi-bin/hgGateway). A program was written to extract the flanking sequences, while masking adjacent regions that were unsuitable for primer location, and to automate the use of *Primer3* [30] for primer sequence design. *Primer3* was unable to design primer sequences to meet the specified parameters for approximately 25% of the loci. The resulting collection consisted of 3113 markers spanning all 38 canine autosomes. Three hundred and eighty-three candidate primer pairs were labeled with M13 tails [31] and screened against individuals from purebred populations to assess polymorphic information content (PIC) across breeds. Candidate primers were evaluated on the Cornell Canine Reference Families (NIH Grant EY006855, G.M. Acland and G.D. Aguirre, Co-P.I.s) to confirm inheritance and establish mutation rates. Markers were further assessed in GeneMapper (Applied Biosystems) for manual edits (ease of scoring) and bin sets that inferred step-wise mutations and minimized the occurrence of microvariant alleles containing insertions or deletions. Primers for the fifteen autosomal loci that were ultimately chosen and for the SRY gene (GenBank Accession number AF107021) were designed using *Primer3* to facilitate their arrangement in the multiplex. Forward primers were labeled with the fluorescent dyes 6-FAM, VIC, NED, and PET for multiplexing. Selected reverse primers were labeled with a seven-base PIGtail [32] to increase specificity or to adjust their position in the multiplex.

2.2. Samples

Samples for validation and databasing consisted of spleen tissue and blood from the Ralston Purina canine repository maintained by Cornell, buccal swabs submitted to the Veterinary Genetics Laboratory for parentage verification or genetic testing, convenience blood samples from veterinary diagnostic laboratories across the U.S., blood and buccal swabs from breed shows, and buccal swabs from shelter dogs. Validation samples were extracted using an organic (phenol–chloroform) extraction protocol [33]. Blood sample buffy coats for population studies were washed followed by extraction with sodium hydroxide and then neutralized with hydrochloric acid and Tris–HCl buffer (French National Institute for Agricultural Research, personal communication, 1995; comparable to the protocol reported by Graffy [34], but without the final concentration and washing steps). DNA quantitation was performed using a TaqMan® assay [35,36] on an Applied Biosystems 7300 Real-Time PCR System (Foster City, CA). Species for which qPCR assays have not been implemented were quantified by spectroscopy on an Eppendorf BioPhotometer according to the manufacturer's instructions.

2.3. PCR amplification

Multiplexed PCR amplification was performed in 25 μL reactions on Applied Biosystems 2720 Thermal Cyclers using 1 μL DNA extract, 0.4× Titanium Taq™ polymerase (BD Biosciences), 1× Titanium Taq™ PCR Buffer (BD Biosciences), 200 μM dNTPs (Invitrogen), 0.09–0.45 μM primer and molecular grade water (Fisher Scientific) to volume. PCR began with a 1 min activation step at 95 °C followed by 31 cycles of 30 s at 95 °C, 30 s at 62 °C, 1 min at 72 °C, and a final extension for 30 min at 72 °C.

2.4. Capillary electrophoresis and data analysis

PCR product was diluted 1:10 in double-distilled water, and 1 μL of that dilution was further diluted into 10 μL HiDi Formamide (Applied Biosystems) and 0.0625 μL GeneScan™ 500 LIZ® size standard (Applied Biosystems). Fragment separation was performed on an Applied Biosystems 3730 DNA Analyzer using the GeneMapper36_POP7 run module. Data were analyzed using both STRand [37] and GeneMapper® (Applied Biosystems) software.

2.5. Construction of an allelic ladder

Samples spanning the allelic range for each locus were assembled from our database collections. When possible, homozygotes were chosen. If the required homozygote was not present in the sample collection, alleles were separated on agarose gel and the desired allele was excised. Target sequences were amplified with the associated primer pair from the multiplex. Five μL of the

resulting product was diluted into 1500 μL of sterile water to make a working stock for each allele. To create the allelic ladder, 5 μL of the working stock for each allele in a locus was combined into a cocktail with an equivalent amount of sterile water. A serial dilution of the allelic ladder cocktail (1:10 to $1:10^6$) was amplified and run under standard conditions to determine the optimum dilution factor and to evaluate the balance between the ladder alleles. Allele stock was added to the cocktail to achieve the desired balance.

2.6. Allele sequencing

Allele-specific repeat sequence data were obtained through direct sequencing of isolated ladder amplicons, amplified sample extracts, or clones. Where feasible, the associated primers from the multiplex were used. For primers that were located close to the repeat region, new primers were designed further away from the repeat region to obtain full sequence reads in both directions that clearly elucidated the repeat motif and flanking sequences. PCR cleanup was performed by incubating 5 μL PCR product with 2.5 units Exonuclease I (USB Corporation, Cleveland, OH), 0.25 units Shrimp Alkaline Phosphatase (SAP) (USB Corporation), and 1 μL SAP 10× reaction buffer (USB Corporation) for 30 min at 37 °C followed by 15 min at 80 °C. Cycle sequencing was performed using BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) with 40 cycles of 20 s @ 95 °C followed by10 s @ 50 °C and 4 min at 60 °C. BigDye was removed using Performa® DTR Gel Filtration Cartridges or 96-well short plates (Edge Biosystems, Gaithersburg, MD) according to the manufacturer's instructions. The resulting sequences were analyzed in Sequencher v. 4.9 (Gene Codes Corporation, Ann Arbor, MI).

2.7. Match probabilities

We calculated locus-specific match probabilities for the casework examples (SWGDAM 2.6) that accounted for population structure among breeds and inbreeding within breeds using the formulae of Balding and Nichols [38] and Ayers and Overall [39], respectively. Values for theta were chosen according to estimates using the markers as described in SWGDAM 2.7. We chose a value for θ at the upper end of the range of reasonable values, in particular, the upper 95% confidence limit, but used the point estimate for f, as suggested by Ayers and Overall [39]. In the absence of evidence for linkage (i.e., gametic disequilibrium beyond that due to population structure, see SWGDAM 2.7), a multilocus genotype match probability was obtained as the product of all locus-specific match probabilities.

3. Developmental validation

3.1. Characterization, SWGDAM 2.1

Inheritance was established through parentage verification of eight multi-generational families comprising 208 individuals (27 founders and 181 progeny) representing 362 meioses. Mutation rates for each locus were estimated by counting the occurrences of discordant repeat copy number transmission from parent to offspring. Loci possessing mutation rates greater than 1% were excluded from consideration for panel inclusion.

Allele sequence polymorphisms were investigated through sequencing of the ladder amplicons (Section 2.6, above) which allowed for repeat-based nomenclature as recommended by the International Society for Forensic Genetics (ISFG) [40] and incorporated the "one-change rule" as described by Butler et al. [41]. Detection was achieved on a capillary electrophoresis unit as detailed in Section 2.4, above. Chromosomal map locations were confirmed by aligning primer sequences against the published dog genome maintained on the UCSC Genome browser, http://genome.ucsc.edu/.

3.2. Species specificity, SWGDAM 2.2

Primer specificity was evaluated on 10, 1, and 0.1 ng of male DNA from the following species within the order *Carnivora*: wolf (*Canis lupus*), dingo (*Canis lupus dingo*), coyote (*Canis latrans*), golden jackal (*Canis aureus*), fox (*Vulpes vulpes*), African wild dog (*Lycaon pictus*), domestic cat (*Felis cattus*), harbor seal (*Phoca vitulina*), and black bear (*Ursus americanus*). Species that may be present in the home as food, pet, or pest were also queried: pig (*Sus scrofa*), cow (*Bos taurus*), sheep (*Ovis aries*), horse (*Equus equus*), and mouse (*Mus musculus*), as well as human (*Homo sapiens*). Bacteria commonly associated with dog biological samples [42]—*Escherichia coli*, *Staphylococcus aureus*, and *Pasteurella aerogenes* (ATCC, Manassas, VA)—were screened against the panel under the same conditions.

3.3. Sensitivity, SWGDAM 2.3

Assay sensitivity was evaluated using nine dog genomic DNA extracts with starting concentrations of between 1.44 and 4.24 ng/μL. Samples were serially diluted in TE buffer down to 0.034–0.067 ng/μL and amplified under standard conditions. All dilutions were quantified by qPCR prior to amplification with the multiplex. Sensitivity was assessed by examining peak height and peak imbalance which, for the purpose of this study, was defined as the peak height of the longer allele in a heterozygous individual exceeding the peak height of the shorter allele.

3.4. Stability, SWGDAM 2.4

The effects of environmental and chemical degradation have been investigated extensively in validation studies on human DNA profiling systems [28], and those effects are expected to be the same for other mammalian species. We elected to perform a degradation study on the effects of enzymatic degradation of DNA over time based on those done by Swango et al. [43]. A series was prepared by adding sterile water and 10× reaction buffer to dog DNA to make 110 μL of a 141.5 ng/μL solution. Ten μL were removed as a negative control, and 2.0 μL DNase I (1 U/μL) (Fermentas Life Sciences, Glen Burnie, MD) were added to the remaining solution. The reaction was incubated at room temperature, and a series was prepared by removing 10 μL volumes at 1, 2, 3, 4, 5, 7.5, 10, 12.5, 15, and 20 minute time points. Each aliquot in the series was stopped by adding 2 μL of 25 mM EDTA and heating at 65 °C for 15 min. One μL of a 1:3 dilution of this degradation series was amplified with the multiplex and run on an AB 3730 to assess panel efficacy on fragmented DNA.

3.5. Reproducibility, SWGDAM 2.5

Primer sets and quantified DNA extracts were sent to the California Department of Justice Bureau of Forensic Services, DNA Laboratory (CAL-DOJ); the University of Queensland Animal Genetics Laboratory, Australia; Genetic Technologies Limited, Fitzroy, Victoria, Australia; and the Veterinary Genetics Laboratory at the University of Pretoria, South Africa. Each participating laboratory evaluated the panel using their standard reagents and equipment, with the exception of the CAL-DOJ who used both Titanium Taq and AmpliTaq Gold® (Applied Biosystems) on an AB 3130xl. The Veterinary Genetics Laboratory at the University of Pretoria, South Africa, used SuperTherm GOLD taq polymerase (Southern Cross Biotechnology) on an AB 3130xl; the University of

Table 1
Locus information for DogFiler multiplex.

Locus name	Repeat location	Fluorophore or PIGtail	Primer sequence	Size range, bp	Tasha's profile repeat no.	Mutation rate
VGL0760	7: 60065445	6-FAM Untailed	F: gcagattcaggacaaagacca R: ggcccagaaaaggataggag	276–340	13	0.0029
VGL0910	9: 10224058	NED GTTTCTT	F: acacttgctcccacgttcct R: accttatgcccaaagcgtgt	282–350	13	0.0054
VGL1063	10: 63191724	PET GTTTCTT	F: agccacagagcctgagagtg R: caatcaccaccttccctcct	86–138	11/12	None found
VGL1165	11: 65356234	VIC GTTTCTT	F: atcttcctctggcaccacct R: ggccctaaatcccatgactg	191–271	15/27	0.0027
VGL1541	15: 41210435	6-FAM GTTTCTT	F: gagctcctgatggaagagctta R: catcctgtccgtgacttcaa	184–240	17/19	0.0054
VGL1606	16: 6468079	PET GTTTCTT	F: agccttcggggtcagatgt R: cccacactgaagctaaactgc	272–340	20	0.0054
VGL1828	18: 28419883	NED GTTTCTT	F: agattgcgcctttggaagt R: cttttggcttcctgctctgt	220–284	19	None found
VGL2009	20: 9290711	PET GTTTCTT	F: ccatttaccagaatttgaagctg R: cccgggaaacttttctgaat	144–184	12/13	0.0027
VGL2136	21: 36673167	VIC Untailed	F: tgccaactgtttttaaaggtaaca R: gcatggagaaaaagcaggtg	91–135	14	None found
VGL2409	24: 9197210	NED GTTTCTT	F: aagcaggtgcttcaacctctg R: aggatagacctccataactgacca	108–156	16/21	0.0027
VGL2918	29: 18216971	PET Untailed	F: gattcttcctggatatgctgcttt R: ggaaasatgtgttttcccttca	188–260	15	None found
VGL3008	30: 8845920	6-FAM Untailed	F: agaacacggttatttgctaggc R: aagagccaacagcagcaga	110–178	18/19	0.0027
VGL3112	31: 12944088	NED gtttctt	F: agccaatagagcattaagtagagctg R: ttgtgtaatgtgtgaatttaagggaat	185–217	16/17	0.0027
VGL3235	32: 35527890	VIC GTTTCTT	F: ggcgactcttctccctttctt R: tctggactgagacagtctgaaaat	267–327	15	None found
VGL3438	34: 38458581	VIC GTTTCTT	F: acgcttgtgggtgctacact R: agcagtgatgagcagagatgg	136–188	14/21[a]	0.0027
SRY	Y	NED	F: gaacgcattcttggtgtggt R: tgatctctgagttttgcatttgg	80	Neg.	None found

[a] The published sequence for locus VGL3438 contained 23 repeats, but cloning and sequencing of TASHA confirmed that she is 14/21 at that locus.

Queensland Animal Genetics Laboratory used a polymerase produced by Bresatec (Adelaide, South Australia) on an AB 3130xl; Genetic Technologies Limited used HotStar Taq DNA Polymerase (Qiagen) on both AB 3130xl and AB 3730xl platforms.

3.6. Case-type samples, SWGDAM 2.6

Retained extracts from adjudicated casework were re-examined using DogFiler. These samples included feces from clothing in a sexual assault, blood swabbed from a dog fighting pit belonging to a professional athlete, saliva from a child's clothing in a fatal dog attack, and shed dog hair recovered from blankets wrapped around a homicide victim. These casework examples are frequently encountered in domestic animal forensic casework and are representative of both the DNA source type and the kinds of crimes associated with domestic dogs.

3.7. Population studies, SWGDAM 2.7

A total of 2234 samples representing 238 breeds and 277 mixed-breed dogs were profiled (Supplemental Table 1). A subset of twenty-five popular breeds was evaluated for Hardy–Weinberg equilibrium, linkage disequilibrium, and coancestry. We tested Hardy–Weinberg equilibrium and estimated coefficients of inbreeding (F_{IS} or f) within breeds, fixation indices (F_{ST} or θ) among breeds, and total inbreeding (F_{IT}) using an analysis of variance framework [38] implemented in Arlequin v3.5 [44]. We tested for gametic (''linkage'') disequilibrium among loci within breeds using Genepop on the Web (v 4.0.10) [45]. Gametic disequilibrium between loci can occur as a consequence both of systematically co-segregating (e.g., physically linked) loci and of inbreeding. Use of the product rule to combine match probabilities across loci requires that all loci segregate independently. Therefore, we assessed whether observed cases of significant ($P < 0.05$) gametic disequilibrium resulted at least partly from systematic co-segregation of particular locus-pairs or could be better explained as a consequence solely of inbreeding (and/or chance type I errors). Specifically, we used a Chi-square goodness of fit test to assess the null hypothesis that the frequency of gametic disequilibrium across breeds conformed to a Poisson (i.e., random) distribution [46].

3.8. Mixture studies, SWGDAM 2.8

While the occurrence of mixture profiles in animal casework is encountered less frequently than in human casework, they are often found in samples obtained from dog fighting pits and in attacks on humans or other animals by a dog pack. For evaluating the effects of DNA mixtures on interpretation, two well-characterized and polymorphic dog genomic DNA samples were combined in the following ratios: 1:1, 1:2, 1:3, 1:4, 1:9, and 1:19. One ng of each mixture DNA was amplified with the DogFiler multiplex.

3.9. Precision and accuracy, SWGDAM 2.9

Precision and accuracy were assessed by collecting fragment size data from twelve injections of the allelic ladder and determining the standard deviation from the mean for each allele.

3.10. PCR-based procedures, SWGDAM 2.10

PCR conditions were optimized using a PTC-200 DNA Engine (MJ Research) to evaluate a range of polymerase concentrations and reaction temperatures. The magnesium chloride concentration was constant at 3.5 mM as a component of the PCR buffer system, and the dNTPs were kept at 0.2 mM for continuity with genotyping

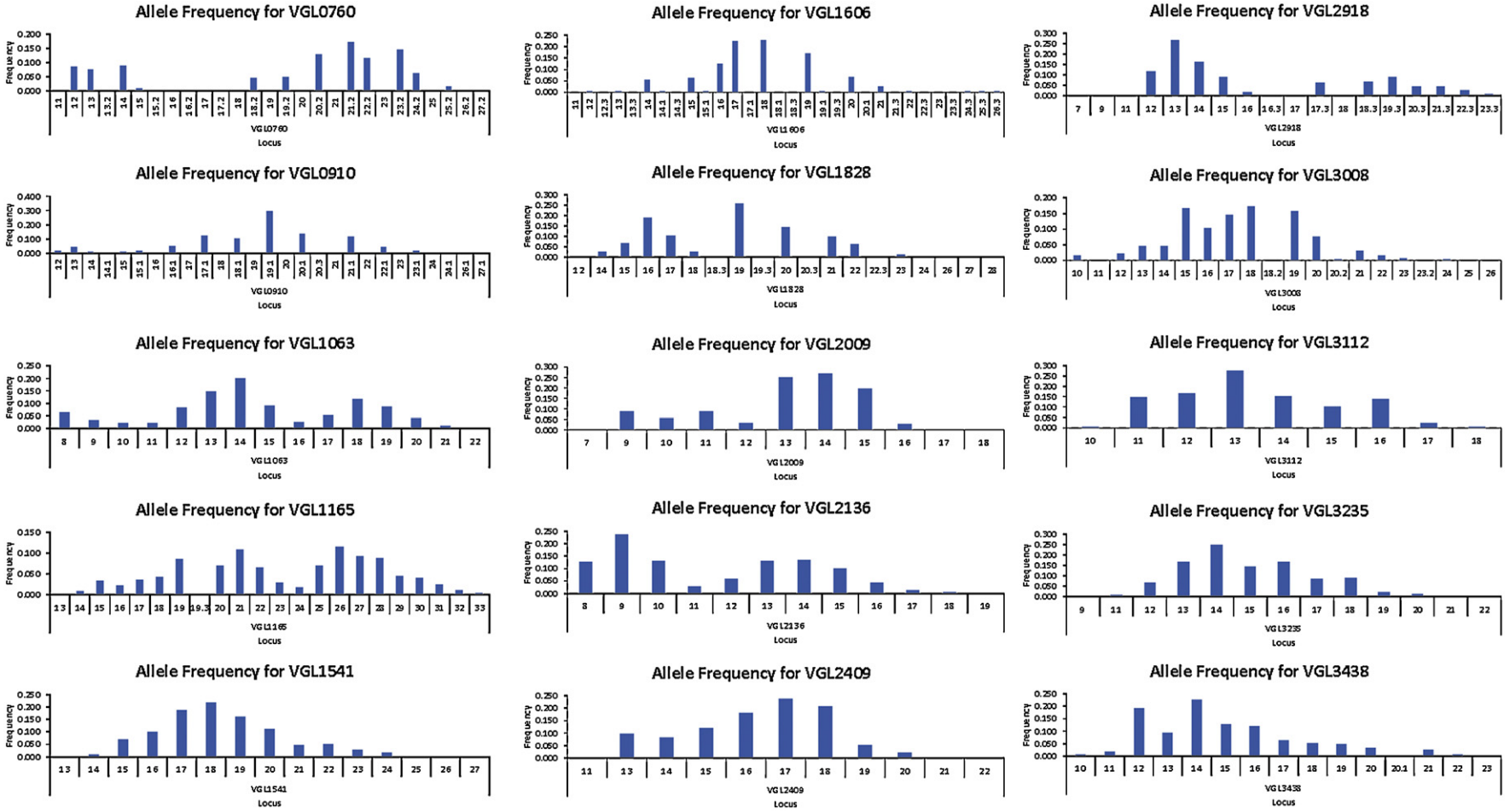

Fig. 1. Allele frequency profiles for the DogFiler loci across 2234 pure bred and mixed-breed dogs.

PCR protocols for other species. The effects of multiplexing (e.g. artifacts) and the potential for intra-locus and inter-locus differential amplification were evaluated through direct observation of database electropherograms. In the absence of a NIST standard, laboratories must select a positive control that is representative of the dogs they test. Gonadal tissue is a readily available source of high-quality material that would otherwise be discarded by spay-neuter clinics. We have chosen a male crossbreed as a positive control. Negative controls include both a negative extraction control and a reagent blank control.

4. Results

4.1. Characterization, SWGDAM 2.1

The chromosomal location for each repeat region, based upon the May 2005 construct of the dog genome assembly, is presented in Table 1 along with the primer sequences, dye labels, fragment size ranges, and mutation rates. For reference, the profile for Tasha, the dog whose genome was sequenced, is also given. All mutations could be accounted for by the insertion or deletion of a single repeat unit from parent to offspring, with insertions outnumbering deletions three to two. When the parental source of discordant transmission could be identified, the mutation was twice as likely to have originated with the dam. This is in contrast to that reported in humans [47–49], although higher mutation rates have been reported to have originated in female as opposed to male barn swallows [50]. The overall mutation rate of 2.4×10^{-3} is virtually identical to the 2.9×10^{-3} reported by Irion et al. [13] for their subset of 66 stable canine microsatellites, and is on the same order of magnitude as that reported for human parentage testing [48,51]. Sequence polymorphisms represented in the allelic ladder are provided in Supplemental Table 2. Three loci contained high frequencies of microvariant alleles-VGL0760, VGL0910, and VGL2918 (Fig. 1). However each of those loci began as a whole repeat, and once an insertion or deletion (indel) occurred, 96–99% of the remaining alleles contained that indel. This allows for consistent bin sets and easy scoring.

4.2. Species specificity, SWGDAM 2.2

Wolves (*C. lupus*), golden jackal (*C. aureus*), and African wild dog (*L. pictus*) produced full 16-locus genotype profiles with some loci exhibiting microvariant and out-of-range alleles when compared to domestic dogs; dingoes (*C. lupus dingo*) produced 15-locus or 16-locus genotype profiles with frequent dropout of locus VGL3438, likely due to a mutation in a primer binding site, and novel microvariant alleles at two additional loci; coyotes (*C. latrans*) produced 15-locus or 16-locus genotype profiles with frequent dropout of locus VGL1063, likely due to a mutation in a primer binding site, and several additional loci exhibiting microvariant or out-of-range alleles when compared to domestic dogs; fox (*V. vulpes*) produced product for all 16 loci with many loci exhibiting microvariant alleles, out-of-range alleles, and fixed alleles or non-STR product; bear (*U. americanus*) was polymorphic for two loci, VGL1063 and VGL0760, as well as amplifying the SRY gene; SRY product and non-specific peaks in VGL1063 and VGL1165 were observed in harbor seal (*P. vitulina*); and no product was observed in the remaining mammalian species. The bacteria tested did not yield product with the exception of a VIC-labeled peak at 131 bp in *E. coli* with a height that directly correlated to the amount of input DNA. While many of the primers amplified across genera within the Canidae family, the presence of product in the non-target species does not invalidate the use of the assay [28].

4.3. Sensitivity, SWGDAM 2.3

All nine samples yielded full 16-locus profiles with template concentrations down to 0.06 ng—below which allelic dropout began to occur. Peak imbalance greater than 10% was observed sporadically in all loci, with 92.5% of the imbalance occurring in template concentrations below 0.25 ng. Locus VGL1828 had the highest frequency of peak imbalance at 21.7% across all runs (template input range from 0.034 ng to 4.24 ng), and the mean occurrence of peak imbalance for all loci was 7% across all runs. The optimum amount of template DNA under these conditions was determined to be 0.5–1.5 ng, although successful genotyping was achieved with a much wider range of input DNA.

4.4. Stability, SWGDAM 2.4

The frequency of obtaining full DNA profiles was inversely proportional to the amount of time the sample was exposed to nuclease activity. As expected, the largest amplicons were the first to fully degrade resulting in a "ski-slope" effect on the electropherograms. At the three-minute time point, more than half of the alleles exhibited dropout. However, amplicons below the 160 bp size standard persisted past the ten-minute time point with one locus still present at the twenty-minute time point. This study demonstrates that exposure of template DNA to environmental degradation by DNases can result in reduced yield of amplified product. To address that problem, we have developed a mini-STR version of DogFiler for use on degraded and inhibited samples (see accompanying submission).

4.5. Reproducibility, SWGDAM 2.5

All laboratories successfully genotyped the samples provided, and all but one obtained full profiles. Allele calls were consistent within laboratories; however, two laboratories reported non-consensus allele calls at two different markers when compared to the reported types. This was prior to the development of an allelic ladder, so the use of a ladder should resolve allele sizing differences between laboratories. Peak heights varied greatly

Table 2
Random match probability estimates for the four casework examples include using the upper 95% confidence estimate for $\theta = 0.142$ and accounting for three levels of inbreeding coefficient (f) using the point estimate (0.054), the lowest estimated for any breed (0.003), and the highest observed in any breed (0.115); estimates based on breed-specific allele frequencies and inbreeding estimates; and, most conservatively, as match probabilities between siblings.

	Formula used	F_{ST}	F_{IS}	Chihuahua Texas	Pit Bull Terrier California	Pit Bull Terrier Virginia	Pomeranian California
Match probability pooled data	Balding and Nichols	0.142		2.25E−15	1.26E−16	7.46E−16	1.98E−18
	Ayers and Overall	0.142	0.054	1.15E−14	4.44E−16	4.26E−15	3.45E−17
	Ayers and Overall	0.142	0.003	3.15E−14	1.22E−15	1.04E−14	8.49E−17
	Ayers and Overall	0.142	0.115	7.82E−15	3.04E−16	2.19E−15	1.80E−17
Match probability within breed	Balding and Nichols	0		6.44E−21	3.59E−24	4.74E−22	3.32E−23
	Ayers and Overall	0	Breed-specific	5.76E−21	3.93E−24	1.35E−22	3.88E−23
Match probability sibling	Evett and Weir			2.84E−07	1.37E−07	2.75E−07	5.39E−08

Table 3
Sample size (n), expected (H_E) and observed (H_O) heterozygosity, and inbreeding coefficient (F_{IS}) estimated for 25 popular breeds.

Breed	n	H_E	H_O	F_{IS}
American Pit Bull Terrier	64	0.839	0.819	0.025
Australian Shepherd	32	0.812	0.748	0.080**
Beagle	30	0.801	0.784	0.021
Border Collie	31	0.807	0.763	0.055**
Boxer	36	0.689	0.641	0.071**
Bulldog	105	0.65	0.611	0.060***
Chihuahua	36	0.849	0.82	0.034
Cocker Spaniel	38	0.71	0.677	0.046*
Collie	30	0.592	0.524	0.115***
Dachshund	30	0.798	0.756	0.055*
Doberman Pinscher	31	0.669	0.596	0.111**
German Shepherd	30	0.752	0.684	0.092**
Golden Retriever	56	0.738	0.72	0.024
Jack Russell Terrier	31	0.837	0.834	0.003
Labrador Retriever	106	0.804	0.776	0.034**
Maltese	31	0.783	0.731	0.067**
Miniature Pinscher	30	0.754	0.684	0.093**
Pomeranian	30	0.79	0.72	0.090***
Poodle	41	0.82	0.735	0.105***
Pug	31	0.641	0.6	0.065*
Schnauzer	33	0.71	0.653	0.083**
Shetland Sheepdog	33	0.71	0.665	0.064**
Shih Tzu	41	0.744	0.735	0.012
Siberian Husky	30	0.754	0.722	0.043
Yorkshire Terrier	42	0.812	0.789	0.029

* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

with different polymerases and equipment platforms. CAL-DOJ reported that the use of AmpliTaq Gold required 6 units of enzyme and 38 PCR cycles to obtain the same results as they obtained using the Titanium Taq protocol and reagents. However, we obtained full profiles (5000 rfu) using 4 units of AmpliTaq Gold and 31 cycles of PCR when run on an AB 3730 unit. The lower signal strength observed by CAL-DOJ was likely a function of reduced sensitivity on the AB 3130xl platform. Likewise, Genetic Technologies found that their polymerase was too robust for use with this panel on the AB 3730xl, and they obtained cleaner traces on the AB 3130xl. After reviewing electropherograms from laboratories around the world, it is evident that this panel works well with a variety of enzymes and equipment platforms. Further optimization specific to each laboratory's polymerase and capillary electrophoresis unit would be required as a part of internal validation.

Table 4
Descriptive statistics of loci based on 2294 dogs from multiple breeds, including number of alleles (Na), probability of identity (P_{ID}), polymorphic information content (PIC), expected heterozygosity (H_E), observed heterozygosity (H_O), power of discrimination (PD), and power of exclusion (PE).

Locus	Na	P_{ID}	PIC	H_E	H_O	PD	PE
VGL0760	27	0.022	0.882	0.892	0.746	0.978	0.478
VGL0910	27	0.037	0.836	0.850	0.717	0.963	0.439
VGL1063	15	0.022	0.881	0.890	0.730	0.978	0.456
VGL1165	22	0.010	0.924	0.928	0.770	0.990	0.513
VGL1541	15	0.035	0.844	0.859	0.739	0.965	0.468
VGL1606	30	0.044	0.823	0.842	0.708	0.956	0.428
VGL1828	19	0.040	0.830	0.847	0.671	0.960	0.383
VGL2009	11	0.062	0.783	0.809	0.657	0.938	0.367
VGL2136	12	0.034	0.846	0.861	0.712	0.966	0.433
VGL2409	11	0.048	0.814	0.835	0.687	0.952	0.402
VGL2918	18	0.034	0.845	0.859	0.740	0.966	0.470
VGL3008	20	0.028	0.863	0.875	0.720	0.972	0.443
VGL3112	9	0.054	0.802	0.825	0.685	0.946	0.400
VGL3235	13	0.043	0.826	0.844	0.642	0.957	0.352
VGL3438	15	0.033	0.848	0.863	0.707	0.967	0.426

4.6. Case-type samples, SWGDAM 2.6

Full DNA profiles were obtained from all samples tested (Supplemental Table 3). The panel proved to be very discriminating with an average heterozygosity of 0.883 for those four samples. The presence of a private allele only found in Pomeranians in our database further supported the prosecutor's allegation that the dog belonging to the suspect's mother was the source of the hair in that case.

Match probability—Estimates of random match probability accounting for population genetic structure were calculated both with and without compensating for within-breed coancestry (i.e., inbreeding; Table 2). The estimates for θ and f used in these calculations were based on data presented below (SWGDAM 2.7). For comparison, θ was set to zero and breed-specific allele frequencies were instead used to calculate random (within-breed) match probabilities. Regardless of whether inbreeding (f) was incorporated [39] or not [38], these breed-specific estimates were far lower than those based on the overall allele frequencies accounting for population structure as per [38] or [39]. Lastly, to provide an extremely conservative reference, sibling match probabilities were calculated, which, as expected, were considerably higher than all other estimates, but, nevertheless sufficiently low as to confidently exclude chance as a reasonable explanation for a match [52].

4.7. Population studies, SWGDAM 2.7

Pooling across all 2234 dogs regardless of breed, we obtained an estimate of overall heterozygosity of 0.859 (SE = 0.007) and individual heterozygosity at 0.709 (SE = 0.006), corresponding to a total inbreeding estimate of $F_{IT} = 0.175$. The estimates based on the subset of 1028 dogs from 25 breeds with $\geq$30 individuals were similar in terms of overall average heterozygosity ($0.856 \pm SE = 0.008$) and individual heterozygosity ($0.712 \pm SE = 0.016$; Table 3), corresponding to a similar estimate of F_{IT} (see below). The heterozygosity within breeds was estimated at 0.755 ($\pm$SE = 0.014). Subsequent analyses were, therefore, based on the subset of dogs from the 25 breeds with $\geq$30 individuals. Allele frequency profiles for the 2234 multi-breed database are presented in Fig. 1. Other basic descriptors of these loci are presented in Table 4.

Most breeds showed deviations from Hardy–Weinberg equilibrium across loci as indicated by inbreeding coefficients, F_{IS}, significantly >0 (Table 3). A total of 240 of 2625 (9.1%) within-breed locus pairs were found to be in gametic disequilibrium, which is greater than the 5% expected by chance (due to type I errors). There was no consistency among locus pairs and breeds, and the distribution did not differ significantly from random (Chi square, 6 df = 5.50, $P = 0.48$), indicating that gametic disequilibrium could be explained by inbreeding alone. The coancestry estimates were as follows: $F_{IT} = 0.171$ (95% CI = 0.151–0.192), $F_{ST} = \theta = 0.124$ (95% CI = 0.107–0.142), and $F_{IS} = f = 0.054$ (95% CI = 0.045–0.063). These coancestry estimates compare to the $F_{IT} = 0.216$, $F_{ST} = 0.106$, and $F_{IS} = 0.123$ reported across breeds in a study of purebred dogs using other STRs [23]. The comparatively higher inbreeding (f) relative to among-breed structure (θ) in their study could reflect higher allelic dropout associated with their markers or greater substructure among dogs used in that study (e.g., due to inclusion of multiple family groups or litters).

4.8. Mixture studies, SWGDAM 2.8

All alleles were clearly discernible in the 1:1, 1:2, and 1:3 mixtures. At a 1:4 mix (0.2 ng minor contributor, 0.8 ng major contributor), some alleles began to become indistinguishable from stutter peaks. The choice of polymerase resulted in increased

stutter product (see PCR-based Procedures, SWGDAM 2.10), so incorporation of a different polymerase may enhance the ability to discern mixture contributors present at less than 1:4. However, even at a 1:19 mixture, 75% of all alleles were unambiguously identified for the two dogs used in this study.

4.9. Precision and accuracy, SWGDAM 2.9

All allele sizes fell within three standard deviations of the mean for repeated injections of the allelic ladder. The allelic ladder contained some fragments that were phenotypically homozygous but were proven to be heterozygous when sequenced. However, this did not affect the fragment size, and the consistency and reproducibility of the allele sizing confirms the reliability of genotyping profiles obtained under this system.

4.10. PCR-based procedures, SWGDAM 2.10

PCR optimization determined that a wide range of PCR conditions yielded reliable product. Adequate product was obtained throughout the annealing temperature range tested (56–64 °C) with the greatest product yield occurring between 61 °C and 64 °C. Annealing times were assessed from 5 s to 60 s with full profiles obtained at 10 s and robust peak heights obtained at 20 s and above. Increased extension time produced an increase in overall peak height with a 50% increase occurring when incorporating a 45 s compared to a 10 s extension. Increased extension time translated to increased product yield for larger fragments resulting in a more balanced multiplex. Increasing the amount of polymerase produced a small increase in peak height, with an average of 7% per 0.1 unit for low-template (0.06 ng) samples.

Occasional intra-locus peak imbalance was observed and was likely due to incomplete primer annealing during PCR. Sequencing of those rare alleles would identify sites that may benefit from primer redesign. A transversion from cytosine to guanine with a frequency of 0.025 was found in the primer-binding region of locus VGL2918, so a degenerate reverse primer was incorporated (Table 1) to promote amplification of the alleles containing the SNP. Optimal amplification was achieved with a two-to-one ratio of degenerate reverse primer to reverse primer incorporating a cytosine at position 6.

Stutter ratios ranged from 5.4% to 25% and were locus-dependent, with the average stutter percentage being 14.8% for Titanium Taq polymerase and 9.4% for AmpliTaq Gold. This difference between polymerases was due to specific loci having as much as a 2.8 fold increase in stutter percentage with Titanium Taq over AmpliTaq Gold, while other loci produced equivalent stutter regardless of the polymerase used. Locus-specific stutter percentage has also been reported for the core human loci with percentages ranging from 6.4% to 13.7% [53].

While our goal was to produce a balanced multiplex, inequities in product yield naturally occur due to degradation, homozygosity, and wide separation of alleles in heterozygotes that are sample dependent. Non-specific artifacts occurred infrequently and were easily detected due to their low peak heights, their peak shape, and their off-ladder sizing. All profiles were unambiguous and easily scored, and in no instance did the presence of an artifact result in an erroneous allele call. The use of the allelic ladder further reinforces confidence in allele calls between runs, across equipment platforms, and between laboratories.

5. Discussion

This validation of the DogFiler panel represents the first published non-human multiplex to fully address all SWGDAM recommendations for panel development. This validation was further enhanced through development of the allelic ladder that revealed a primer binding mutation promoting spurious null alleles at one locus which was consequently resolved through primer redesign. The large number and variety of breeds represented in the population studies described in this study elucidated potential multiplexing issues (e.g. range overlap, high mutation rates, and microvariant alleles) that have compromised other canine STR panels [25]. The opportunity to select new markers specifically for forensic genetics enabled us to create a de novo panel for canine forensic testing with the increased stringency essential for profiling populations such as domestic dogs where high levels of inbreeding and consequent loss of diversity exist.

The high level of population structuring in dogs can be attributed to both their evolutionary history and, in particular, recent breeding practices. There is evidence for the relatively recent derivation of dogs from Eurasian wolves (~16,000 ya) [14,54] followed by a rapid phenotypic diversification. This wide phenotypic variation has fostered creation of a diverse array of breeds through deliberate selection for and concentration of desirable traits. Over 1000 breeds have been reported historically [55] and approximately 500 internationally recognized breeds exist today. The majority of those breeds were developed in the last 200–400 years through intense selection for morphological and behavioral traits, targeted introgression for reasons of health or conformance, and genetic bottlenecks due to disease, breed popularity, or economics. Along with increased diversity *among* breeds has come a loss of genetic diversity *within* breeds. Restrictions imposed by breed registries continue to further reduce genetic variability and to increase population structure [56], with the magnitude of substructure in some breeds approximating the divergence found among breeds [57]. Separate stocks maintained by registered breeders versus "backyard breeders" who draw from different pools within the same breed could further contribute to substructure.

When calculating match probabilities in forensic casework, it is imperative to take a conservative approach that does not overstate the power of a DNA match. The substantial population structure present in domestic dogs poses a challenge for forensic applications. Random match probabilities are highly sensitive to structure and will tend to be substantially underestimated when structure and substructure are not accounted for. In principle, if the breed and corresponding breed-specific allele frequencies are known, this can be done directly. However, in practice, there are obstacles to such an approach. For example, breed identification depends upon the knowledge and experience of the police officer, veterinarian, animal control officer, or shelter worker involved in each case. This is especially problematic when dealing with mixed-breed dogs which make up over 40% of the U.S. dog population [1], making the application of a breed-specific database even more challenging. Population structure has been addressed in both human and dog DNA cases by utilizing allele frequencies obtained from an admixture of breeds or populations and adjusting by an overall estimate of population structure [7,52]. Here we demonstrated that such estimates can be extremely conservative and therefore valid in a courtroom setting. For example, in the four cases we presented, match probabilities based on overall allele frequencies, accounting for structure and/or inbreeding, were >7 orders of magnitude higher than those estimated using breed-specific allele frequencies and this remained the case even when inbreeding was accounted for in the latter estimates. In cases where there is reason to believe that a dog could be mistaken for its sibling, then it would appear justified to use match probabilities of siblings as proposed by Ogden et al. [58] which are much more conservative than other estimates. Even in such cases, however,

the polymorphism of the DogFiler panel was sufficient to maintain match probabilities for siblings on the order of 10^{-7} (Table 2).

5.1. Potential problems

Due to the close positioning of the ranges for each marker in the multiplex, new alleles may be found that fall into the range of an adjacent marker with the same dye label. While primer redesign can provide easier identification of those alleles, the rarity of those alleles would likely make that option impractical. If an out-of-range allele is suspected, amplification using individual primer pairs or miniplexes would resolve that ambiguity. Individual primer pairs were also used to confirm an observation of three alleles at locus VGL2918 in one dog.

6. Conclusions

While there are many markers available for individual canine DNA profiling, no published assemblage has proven to be as robust, discriminating, and accurate as the DogFiler panel. It has demonstrated utility as a tool for both forensic canine DNA analysis as well as a supplemental panel for parentage verification in highly inbred matings due to the high PIC values for the loci (mean >0.84). This panel has been accepted in court in multiple states and forms the core of the dog fighting database established through the joint efforts of the ASPCA, the Missouri Humane Society, the Louisiana SPCA, and the UC Davis Veterinary Genetics Laboratory. Reduced-amplicon mini-STR versions of these loci have undergone validation and are now being successfully applied to degraded or inhibited samples in criminal casework. The combination of the DogFiler multiplex, a set of canine miniSTRs (accompanying submission), and the allelic ladder represents the first animal DNA testing procedures to fully parallel the human forensic model. The implementation of this panel in crime laboratories has the potential to greatly expand the probative value of crime-scene canine biological evidence.

Acknowledgements

This effort was funded in full by the Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California-Davis. We gratefully acknowledge the following contributors for samples used in this project: Purina and Dr. Greg Acland, Cornell University; The Center for Companion Animal Health, University of California, Davis; and Charles Clark, IDEXX Laboratories, Inc. The authors also thank the California Department of Justice, Bureau of Forensic Services Laboratory, Richmond, California; the Veterinary Genetics Laboratory at the University of Pretoria, South Africa; and Genetic Technologies Limited, Fitzroy, Victoria, Australia, for their evaluation and feedback on the panel. We are deeply indebted to Alison Ruhe for conception of this project and her commitment to it, to Aaron Wong for developing the software to scan the canine genome, and to Dr. John Butler of NIST for his advice on repeat motif nomenclature.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2012.07.001.

References

[1] APPMA National Pet Owners Survey, Greenwich, CT, 2006.
[2] F. D'Andrea, F.R.C. Fridez, Preliminary experiments on the transfer of animal hair during simulated criminal behavior, J. Forensic Sci. 43 (6) (1998) 1257–1258.
[3] C. Tarditi, R. Grahn, J. Evans, J. Kurishima, L.A. Lyons, DNA sequencing of cat hair: an informative forensic tool, J. Forensic Sci. 56 (Suppl. 1) (2011) S36–S46.
[4] M.A. Menotti-Raymond, V.A. David, S.J. O'Brien, Pet cat hair implicates murder suspect, Nature 386 (1997) 774.
[5] S. Müller, G. Flekna, M. Müller, G. Brem, Use of canine microsatellite polymorphisms in forensic examinations, J. Hered. 90 (1) (1999) 55–56.
[6] G.G. Shutler, P. Gagnon, G. Verret, H. Kalyn, S. Korkosh, E. Johnston, J. Halverson, Removal of a PCR inhibitor and resolution of DNA STR types in mixed human-canine stains from a five year old case, J. Forensic Sci. 44 (1999) 623–626.
[7] J. Halverson, C. Basten, Forensic DNA identification of animal-derived trace evidence: tools for linking victims and suspects, Croat. Med. J. 46 (4) (2005) 598–605.
[8] P.M. Schneider, Y. Seo, C. Rittner, Forensic mtDNA hair analysis excludes a dog from having caused a traffic accident, Int. J. Legal Med. 112 (1999) 315–316.
[9] Z. Padar, et al., Canine microsatellite polymorphisms as the resolution of an illegal animal death case in a Hungarian zoological gardens, Int. J. Legal Med. 115 (2) (2001) 79–81.
[10] M. Binns, N. Holmes, E. Marti, N. Bowen, Dog parentage testing using canine microsatellites, J. Small Anim. Pract. 36 (11) (1995) 493–497.
[11] S. Denise, et al., Power of exclusion for parentage verification and probability of match for identity in American Kennel Club breeds using 17 canine microsatellite markers, Anim. Genet. 35 (1) (2004) 14–17.
[12] M. Fredholm, A. Winterø, Variation of short tandem repeats within and between species belonging to the Canidae family, Mamm. Genome 6 (1) (1995) 11–18.
[13] D. Irion, A. Schaffer, T. Famula, M. Eggleston, S. Hughes, N.C. Pedersen, Analysis of genetic variation on 28 dog breed populations with 100 microsatellite markers, J. Hered. 94 (1) (2003) 81–87.
[14] J.F. Pang, C. Kluetsch, X. Zou, L. Luo, H. Angleby, A. Ardalan, C. Ekström, A. Sköllermo, J. Lundeberg, S. Matsumura, T. Leitner, Y. Zhang, P. Savolainen, mtDNA indicates a single origin for dogs south of the Yangtzee River, less than 16,300 years ago, from numerous wolves, Mol. Biol. Evol. 26 (2009) 2849–2864.
[15] E.A. Ostrander, G.F. Sprague Jr., J. Rine, Identification and characterization of dinucleotide repeat (CA)n markers for genetic mapping in dog, Genomics 16 (1) (1993) 207–213.
[16] L.V. Francisco, et al., A class of highly polymorphic tetranucleotide repeats for canine genetic mapping, Mamm. Genome 7 (5) (1996) 359–362.
[17] M.W. Neff, et al., A second-generation genetic linkage map of the domestic dog, *Canis familiaris*, Genetics 151 (2) (1999) 803–820.
[18] A.K. Wong, A.L. Ruhe, B.L. Dumont, K.R. Robertson, G. Guerrero, S.M. Shull, J.S. Ziegle, L.V. Millon, B.A. Payseur, K.E. Broman, M.W. Neff, A canine linkage map constructed with a post-genome logic, Genetics 184 (2) (2010) 595–605.
[19] P. Gill, R. Sparkes, L. Fereday, D. Werrett, Report of the European network of forensic science institutes (ENSFI): formulation and testing of principles to evaluate STR multiplexes, Forensic Sci. Int. 108 (2000) 1–29.
[20] ISAG-International Society for Animal Genetics, 2005 ISAG Panel DOG 0505, www.isag.us/Docs/2005ISAGPanelDOG.pdf.
[21] A. Carracedo, M.V. Lareu, Development of new STRs for forensic casework: criteria for selection, sequencing & population data and forensic validation, in: Proceedings from the Ninth International Symposium on Human Identification, 1998, pp. 89–107.
[22] C. Eichman, B. Berger, W. Parson, A proposed nomenclature for 15 canine-specific polymorphic STR loci for forensic purposes, Int. J. Legal Med. 118 (2004) 249–266.
[23] J. Halverson, C. Basten, A PCR multiplex and database for forensic DNA identification of dogs, J. Forensic Sci. 50 (2005) 352–363.
[24] B. van Asch, A new autosomal STR multiplex for canine genotyping, Forensic Sci. Int.: Genet. Suppl. Ser. 1 (1) (2008) 628–629.
[25] M. Dayton, M.T. Koskinen, B.K. Tom, A.M. Mattila, E. Johnston, J. Halverson, et al., Developmental validation of short tandem repeat reagent kit for forensic DNA profiling of canine biological material, Croat. Med. J. 50 (3) (2009) 268–285.
[26] K. Lindblad-Toh, et al., Genome sequence, comparative analysis and haplotype structure of the domestic dog, Nature 438 (2005) 803–813.
[27] V. Meyers-Wallen, D. Schlafer, I. Barr, R. Lovell-Badge, A. Keyner, Sry-negative XX sex reversal in purebred dogs, Mol. Reprod. Dev. 53 (1999) 266–273.
[28] Scientific Working Group on DNA Analysis Methods (SWGDAM), Revised validation guidelines, Forensic Sci. Commun. 6 (2004) http://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/july2004/standards/2004_03_standards02.htm/.
[29] B. Budowle, P. Garafano, A. Hellman, M. Ketchum, S. Kanthaswamy, W. Parson, W. van Heringen, S. Fain, T. Broad, Recommendations for animal DNA forensic and identity testing, Int. J. Legal Med. 119 (2005) 295–302.
[30] S. Rosen, H.J. Skaletsky, Primer 3, Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html, 1998.
[31] W.S. Oetting, H.K. Lee, D.J. Flanders, G.L. Weisner, T.A. Sellers, R.A. King, Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers, Genomics 30 (1995) 450–458.
[32] M.J. Brownstein, J.D. Carpten, J.R. Smith, Modulation of non-templates nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping, Biotechniques 20 (1996) 1004–1010.
[33] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989
[34] E. Graffy, D. Foran, A simplified method for mitochondrial DNA extraction from head hair shafts, J. Forensic Sci. 50 (5) (2005) 1119–1122.
[35] J. Evans, E. Wictum, M.C. Penedo, S. Kanthaswamy, Real-time polymerase chain reaction quantification of canine DNA, J. Forensic Sci. 52 (1) (2007) 93–96.
[36] C. Lindquist, J. Evans, E. Wictum, Developmental validation of feline, bovine, equine and cervid quantitative PCR assays, J. Forensic Sci. 56 (2011) S29–S35.
[37] R. Toonen, S. Hughes, Increased throughput for fragment analysis on an ABI Prism® 377 Automated Sequencer using a membrane comb and STRand software, Biotechniques 31 (2001) 1320–1324.

[38] D.J. Balding, R.A. Nichols, DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands, Forensic Sci. Int. 64 (1994) 125–140.
[39] K.L. Ayers, D.J. Overall, Allowing for within-subpopulation inbreeding in forensic match probabilities, Forensic Sci. Int. 103 (1999) 207–216.
[40] W. Bar, B. Brinkmann, B. Budowle, A. Carracedo, P. Gill, P. Lincoln, W. Mayr, B. Olaisen, DNA recommendations: further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. International Society for Forensic Haemogenetics, Int. J. Legal Med. 110 (1997) 175–176.
[41] J. Butler, M. Kline, A. Decker, Addressing Y-chromosome short tandem repeat allele nomenclature, J. Gen. Geneal. 4 (2008) 125–148.
[42] D.R. Elliott, M. Wilson, C.M.F. Buckley, D.A. Spratt, Cultivable oral microbiota of domestic dogs, J. Clin. Microbiol. 43 (11) (2005) 5470–5476.
[43] K.L. Swango, W.R. Hudlow, M.D. Timkin, M.R. Buoncristiani, Developmental validation of a multiplex qPCR assay for assessing the quantity and quality of nuclear DNA in forensic samples, Forensic Sci. Int. 170 (2006) 35–45.
[44] L. Excoffier, H.E. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol. Ecol. Resour. 10 (2010) 564–567.
[45] F. Rousset, Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux, Mol. Ecol. Resour. 8 (2008) 103–106.
[46] J.H. Zar, Biostatistical Analysis, 5th ed., Prentice Hall, Upper Saddle River, NJ, 2010
[47] B. Brinkman, M. Klintschar, F. Neuhuber, J. Huhne, B. Rolf, Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat, Am. J. Hum. Genet. 62 (1998) 1408–1415.
[48] H. Gaeda, L. Viriato, C. Vieira-Silva, C. Cruz, I. Lucas, T. Ribeiro, R. Espinheira, STR mutations in paternity investigations: a study of 1-year consecutive cases, Int. Congr. Ser. 1239 (2003) 657–660.
[49] F. Aşicioğlu, F. Oguz-Savran, U. Ozbek, Mutation rate at commonly used forensic STR loci: paternity testing experience, Dis. Mark. 20 (2004) 313–315.
[50] J. Brohede, A.P. Møller, H. Ellegren, Individual variation in microsatellite mutation rate in barn swallows, Mutat. Res. 545 (2004) 73–80.
[51] J.M. Butler, Genetics and genomics of core short tandem repeat loci used in human identity testing, J. Forensic Sci. 51 (2) (2006) 253–265.
[52] I.W. Evett, B.S. Weir, Interpreting DNA Evidence, Sinauer Associates, Inc., Sunderland, MA, 1998.
[53] Life Technologies Corporation, AmpFlSTR® Identifiler® Plus PCR Amplification Kit User's Guide, Part Number 4440211 Rev. D, Foster City, CA, 2012.
[54] B.M. vonHoldt, J.P. Pollinger, K.E. Lohmueller, E. Han, H.G. Parker, et al., Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication, Nature 464 (2010) 898–902.
[55] D.M. Morris, Dogs: The Ultimate Dictionary of Over 1,000 Dog Breeds, Trafalgar Square Publishing, North Pomfret, VT, 2002.
[56] F.C.F. Calboli, J. Sampson, N. Fretwell, D.J. Balding, Population structure and inbreeding from pedigree analysis of purebred dogs, Genetics 179 (2008) 593–601.
[57] S. Björnerfeldt, F. Hailer, M. Nord, C. Vilà, Assortative mating and fragmentation within dog breeds, BMC Evol. Biol. 8 (2008).
[58] R. Ogden, R. Mellanby, D. Clements, A. Gow, R. Powell, R. McEwing, Genetic data from 15 STR loci for forensic individual identification and parentage analyses in UK domestic dogs, Forensic Sci. Int. Genet. 6 (2012) e63–e65.